

6 Strukturierte Interviews und Fragebögen

6.1 Quantitative Methoden in der Ethnologie

In: Bettine Beer / Anika König (Eds.):
Methoden ethnologischer Feldforschung.
Berlin, Reimer, 2020.

Viele Ethnologinnen und Ethnologen verstehen ihre Wissenschaft als eine Disziplin, die überwiegend mit qualitativen Methoden arbeitet. Teilnehmende Beobachtung und offene Interviews werden als die Standardmethoden der Ethnologie angesehen. Strukturierte Interviews, die mit Fragebögen erhoben und quantitativ ausgewertet werden, gelten dagegen als Methoden der Soziologie und als ‚unethnologisch‘. Der Gegensatz von qualitativen und quantitativen Methoden hat oft eine schon fast identitätsstiftende Funktion für EthnologInnen. Zum Teil werden auch bestimmte theoretische Richtungen der Ethnologie auf den einen oder den anderen Methodenbereich festgelegt. Dann heißt es etwa, analytische Ethnologie arbeite quantitativ, interpretative Ethnologie verwende dagegen qualitative Methoden.

Dieser strikte Gegensatz bzw. die hinter diesem Schematismus oft verborgene Ablehnung quantitativer Methoden ist jedoch zu kurz gedacht (Schweizer 1998, Sobo und de Munck 1998). EthnologInnen arbeiten sowohl qualitativ als auch quantitativ. Bernard hält fest: „No method of data collection is perfect. Unstructured interviews and questionnaires produce different *kinds* of data, and it is up to you to decide which method, or combination of methods, is best“ (Bernard 1995: 287, Betonung im Original).

Tatsächlich kommt keine ethnologische Studie ohne Quantifizierungen aus. Quantitative Methoden haben mit Häufigkeits- und Mengenangaben zu tun. Wenn jede Ethnographie auch quantitative Aspekte hat, bedeutet das natürlich nicht, dass ethnographische Monographien stets ausgefüllte Tabellen und Statistiken enthalten. Sehr oft bleiben Quantifizierungen implizit, vage und damit ungenau. Es heißt dann etwa über die Nuer: „Families often change their place of residence from one part of a village to another [...]“ (Evans-Pritchard 1969: 65, meine Betonung). Oder, zu Beginn dieses Beitrags: „Viele Ethnologinnen und Ethnologen ...“ Wir erfahren nicht genau, wie oft Nuer umziehen oder für wie viele EthnologInnen tatsächlich diese Aussage gilt. Eine genaue Quantifizierung ist vielleicht auch gar nicht nötig – es kommt eben vor allem darauf an, auszudrücken, dass Nuer überhaupt immer wieder umziehen, dass sie nicht dauerhaft am selben Ort wohnen. Aber manchmal wäre eine genauere Zahlenangabe doch sinnvoll. Dies ist vor allem dann der Fall, wenn aus der Zahl der Umzüge weitere Schlüsse

gezogen werden sollen oder wenn Vergleiche angestellt werden zwischen Haushalten, die oft umziehen und solchen, die sesshafter sind.

Quantitative Methoden haben zum Ziel, exakte Mengenangaben zu erreichen. Sie implizieren Zählungen und Messungen. Damit diese Messungen gültig sind, müssen sie mit verlässlichen Messinstrumenten erhoben werden. Standardisierte Erhebungsinstrumente sind erforderlich, die sicherstellen, dass allen Befragten dieselben Fragen gestellt werden, sodass die Antworten auch tatsächlich ausgezählt, verglichen und zueinander in Beziehung gesetzt werden können. Wenn Interviewdaten quantitativ ausgewertet werden sollen, muss also mit strukturierten Interviews gearbeitet werden, die in der Regel mit Hilfe von Fragebögen geführt werden. Bei offenen oder unstrukturierten Interviews geben der Ethnologe oder die Ethnologin häufig den Gesprächsfaden zu einem gewissen Grad aus der Hand und lassen ihre GesprächspartnerInnen (mit-)bestimmen, über welche Themen gesprochen wird (siehe dazu auch den Beitrag von Schlehe in diesem Band). Bei strukturierten Interviews ist dagegen der Ablauf des Gesprächs vorgegeben. D. h., sowohl der Wortlaut der Fragen als auch ihre Reihenfolge wird von den Forschenden festgelegt und bleibt für alle GesprächspartnerInnen, die bei der Erhebung befragt werden, gleich. Diese Strukturierung und Standardisierung ist erforderlich, um valide Quantifizierung zu ermöglichen.

Die wichtigste Voraussetzung fürs Zählen und Messen sind definierte Einheiten. Die Definition solcher Einheiten ist in der Ethnologie häufig sehr schwierig. Damit werde ich mich im folgenden Abschnitt auseinandersetzen. Im Weiteren behandelt dieses Kapitel die Standardisierung von Erhebungen, die Konstruktion von Fragebögen, die Auswahl von InformantInnen (das Sampling) und verschiedene Erhebungswege. Mein Beitrag schließt mit einem kurzen Einblick in die Datenauswertung.

6.2 Zählen: Das Problem der Einheiten

Zählen ist einfach, man lernt es spätestens in der Grundschule. Schwierig wird Zählen in der Ethnologie dadurch, dass wir etwas zählen und uns sehr genau darüber klarwerden müssen, was wir da zählen. Wir haben es eben nicht – wie in der Mathematik – mit abstrakten Zahlen zu tun, die für nichts als für sich selbst stehen, sondern mit konkreten Dingen, Personen oder Ereignissen, deren Anzahl uns interessiert. Wir wollen etwas zählen, und dazu muss zunächst definiert werden, was dieses Etwas ist. Wir stehen also vor der Frage der Einheiten des Zählens. Auch das erscheint zunächst vielleicht gar nicht problematisch. Wenn wir wissen wollen, wie viele Menschen in einem Dorf leben, dann zählen wir diese Menschen. Die Einheit ist definiert als

das menschliche Individuum. Aber bei genauerem Nachdenken wird klar, dass diese Definition ungenau und keineswegs eindeutig ist. Denn welche Menschen leben in unserem Dorf? Wie lange etwa muss man sich in dem Dorf aufhalten, damit man zur Bevölkerung gezählt wird? Gehört ein Kind dazu, das ein halbes Jahr bei seinen Großeltern im Dorf lebt, ansonsten aber mit seinen Eltern in der Stadt wohnt? Was ist mit MigrantInnen, die sich mal im Dorf aufhalten, mal woanders? Um eine genauere Definition dessen, was gezählt werden soll, kommt man also nicht herum. Das Ziel der Definition ist eine klare und eindeutige Abgrenzung dessen, was unter eine Kategorie fällt, von dem, was nicht dazu gerechnet wird. Hier zwingt die quantitative Methode also dazu, Kategorien genau zu bestimmen.

Die Definition einer Einheit ist zunächst ein willkürlicher Akt. Ich könnte beispielsweise festlegen: Ein Bewohner des Dorfes ist ein Mensch, der sich mindestens die Hälfte des Jahres in diesem Dorf aufhält. Vielleicht habe ich Aufzeichnungen darüber, wer sich wie lange im Dorf aufhält, dann kann ich objektiv sagen, wer unter diese Definition fällt und wer nicht (vermutlich habe ich solche Aufzeichnungen aber nicht oder nur für einen kurzen Zeitraum; dann ist es schwierig, diese Definition anzuwenden).

Trotzdem kann es sein, dass diese Definition nicht besonders sinnvoll ist. Vielleicht fallen Individuen heraus, die nicht mindestens die Hälfte des Jahres im Dorf leben, die aber im sozialen Gefüge des Dorfes eine wichtige Rolle spielen und die von den anderen DorfbewohnerInnen als EinwohnerInnen betrachtet werden. Was ist dann wichtiger, unsere ‚objektive‘ Definition oder die Ansicht der DorfbewohnerInnen? Wir müssen uns zwischen einer *emischen* und einer *etischen* Definition entscheiden. Über diese beiden aus der Linguistik übernommenen Konzepte hat es in der Ethnologie eine langandauernde Debatte gegeben (Headland et al. 1990). Es gibt keine letztendlich allgemein akzeptierte Definition. Grob gesagt hat sich aber das Verständnis durchgesetzt, dass „emisch“ sich auf das Bedeutungsuniversum der jeweils Untersuchten bezieht, während „etisch“ Kategorien und Definitionen meint, die von den Forschenden von außen an den Untersuchungsgegenstand herangetragen werden. Meine selbst festgelegte Definition wäre also etisch, die Definition von „Bewohner“ durch die EinwohnerInnen des Dorfes selbst wäre dagegen eine emische Definition.

Die Unterscheidung von emischer und etischer Perspektive erleichtert es, die Problematik der Definition von Einheiten der Zählung zu veranschaulichen und zu verstehen, aber sie löst diese Schwierigkeit nicht. Häufig gibt es nämlich nicht ‚die‘ emische Perspektive in einer untersuchten Gruppe, sondern verschiedene, miteinander in Konflikt stehende Sichtweisen. Neben verschiedenen lokalen Ansichten mag es auch noch eine offizielle Definition des „Einwohners“ bzw. der „Einwohnerin“ durch die zuständige staatliche Verwaltung geben, die sich von den lokalen Definitionen wieder

unterscheiden kann. Bei meiner Feldforschung in Gilgit, Nordpakistan, stellte sich beispielsweise heraus, dass es höchst umstritten war, wer im Ort als EinwohnerIn galt und wer nicht. Das Konzept des „Einwohners“ bzw. der „Einwohnerin“ umfasste nämlich eine Reihe von Rechten auf Nutzung knapper Ressourcen (vor allem Land und Wasser), die keineswegs allen BewohnerInnen zugestanden wurden. Aber es gab keine Einigkeit darüber, wer nun Einwohner oder Einwohnerin war und wer nicht. Manche Familien, die von einigen als EinwohnerInnen anerkannt wurden, wurden von anderen als „Leute von außen“ kategorisiert, die keine EinwohnerInnen sind und keine Rechte haben. Der Rückgriff auf die offizielle Definition der Verwaltung brachte hier keine Lösung, denn manche EinwohnerInnen warfen der Verwaltung Korruption vor: Sie habe den Status des Einwohners bzw. der EinwohnerIn auch „Leuten von außen“ eingeräumt, die kein Recht darauf hätten. Man muss sich darüber im Klaren sein, dass in einer solchen Situation der einfache Forschungsakt des Zählens ein **politischer** Akt sein kann, der die Sichtweise einer bestimmten Fraktion in der untersuchten Gruppe stützt und selbst zum Element des Konfliktes werden kann. Nicht umsonst sind Zensuserhebungen und die in ihnen verwendeten Kategorien manchmal extrem umstritten.

Aber die Definition von Kategorien der Erhebung kann auch in weniger konfliktträchtigen Fällen problematisch sein. Bei einer Untersuchung der Berufswege von AbsolventInnen des Hamburger Instituts für Ethnologie wollten wir auch wissen, ob die Befragten zur Zeit der Befragung arbeitslos waren oder nicht. Wir hatten dies zunächst für eine einfache Frage gehalten, mussten aber feststellen, dass „Arbeitslosigkeit“ ebenfalls ein sehr unterschiedlich verstandenes Phänomen ist. Da gibt es zunächst die offizielle Kategorie, die alle umfasst, die beim Arbeitsamt als arbeitslos gemeldet sind. Das Arbeitsamt hat jedoch noch eine zweite Kategorie, die der „Arbeits-suchenden“, die oft nicht weniger arbeitslos sind. Aber nicht auf alle arbeitslosen oder arbeitssuchenden Personen trifft zu, dass sie tatsächlich nicht arbeiten. Einige der Befragten waren gerade dabei, ein Promotionsprojekt vorzubereiten und sahen sich nicht als arbeitslos an, obwohl sie (noch) kein Geld verdienten und beim Arbeitsamt gemeldet waren. Andere betrachteten sich durchaus als arbeitslos, waren aber nicht gemeldet, weil sie keine Leistungen vom Arbeitsamt zu erwarten hatten. Wieder andere lebten von Gelegenheitsjobs, und für sie wechselte der Status beinahe wöchentlich, sodass sie sich selbst nicht mehr sicher waren, ob sie nun arbeitslos waren oder nicht (vgl. Luge-Erhardt et al. 2000).

Ein drittes Beispiel für die Problematik der Definition von Einheiten der Befragung: Bei ihrer Untersuchung in Dörfern im nordpakistanischen Bagrot wollte Monika Schneid (persönliche Mitteilung) auch Daten über den Viehbesitz der Haushalte sammeln. Sie fragte jeweils, wie viele Kühe und

Ziegen ein Haushalt besaß. Erst im Nachhinein stellte sie fest, dass sie mit dieser Frage nicht einfach den Viehbestand der Dörfer erhoben hatte. Vor allem Kühe hatte sie mehr gezählt, als tatsächlich vorhanden waren. Denn die Dorfbewohner praktizierten besonders bei Kühen, die eine teure Investition darstellten, kollektiven Viehbesitz. Dieses *cow-sharing* bedeutete, dass eine Kuh mehrere Besitzer haben konnte. Bei der Befragung hatte aber jeder der (Teil-)Besitzer die Kuh als seinen Besitz genannt, sodass Tiere zum Teil zunächst mehrfach gezählt worden waren.

Wie viele EinwohnerInnen ein Dorf hat, wie viele AbsolventInnen eines Studiengangs arbeitslos sind oder wie viele Kühe jemand besitzt, sind also nicht unbedingt Fragen, die sich ‚objektiv‘ beantworten lassen. Die Beantwortung dieser Fragen ergibt sich nicht einfach ‚aus der Empirie‘. Was die jeweilige Empirie ist, hängt davon ab, wie wir die Fragen formulieren und wie sie von unseren GesprächspartnerInnen verstanden werden. So ergeben sich je nach Definition der Kategorie „EinwohnerInnen“ oder „arbeitslos“ ganz unterschiedliche Zahlen. Wenn man in die Tiefe geht, wird man für sehr viele auf den ersten Blick einfache Fragen feststellen, dass sie nicht so einfach zu beantworten sind, wie sie ursprünglich zu sein scheinen, bzw. dass in ihre Beantwortung zahlreiche Vorannahmen fließen, die nicht unmittelbar offensichtlich sind. Wenn man die Befragung aufgrund dieser Schwierigkeiten nicht gleich wieder aufgeben will, dann muss man sich pragmatisch für eine Definition der jeweils in Frage stehenden Kategorie entscheiden. Welche Definition sinnvoll ist, muss aus dem Forschungskontext heraus entschieden werden, dafür kann es keine allgemeine Lösung geben.

6.3 Die Standardisierung der Erhebung

Aus der Diskussion um die Definition der Einheiten ergibt sich die entscheidende Forderung: Wie auch immer die jeweilige Einheit definiert wird, diese Definition muss für die ganze Erhebung durchgehalten werden. Wir können nicht die Arbeitslosigkeit bei Informant A nach seiner Selbsteinschätzung feststellen, bei Informantin B aber danach, ob sie beim Arbeitsamt gemeldet ist. Die Verwendung verschiedener Definitionen in ein und derselben Untersuchung führt zu inkonsistenten Daten, die nicht untereinander vergleichbar sind und aus denen daher keine gültigen Schlüsse gezogen werden können.

Die Standardisierung der Erhebung dient dazu, die Vergleichbarkeit der erhobenen Daten herzustellen. So soll ein Fragebogen sicherstellen, dass allen InformantInnen dieselben Fragen und auf dieselbe Art und Weise gestellt werden. Der zweite Aspekt ist natürlich nicht vollständig zu erfüllen, denn die Art und Weise des Fragens kann auch von so unkontrollierbaren

Bedingungen wie dem Wetter oder der momentanen psychischen Verfassung des/der InterviewerIn beeinflusst werden. Die Bedeutung einer Frage ergibt sich schließlich nicht nur aus ihrem Wortlaut, sondern etwa auch aus ihrer Betonung. Tatsächlich kann ein Fragebogen also nur gewährleisten, dass die Unterschiede der Art und Weise, wie dieselben Fragen verschiedenen GesprächspartnerInnen gestellt werden, minimiert werden.

Das strukturierte Interview ist eindeutig von der oder dem Forschenden dominiert. Die jeweiligen GesprächspartnerInnen haben auf die Themen, die besprochen werden, keinen Einfluss. Sie können lediglich die Beantwortung von Fragen verweigern. Dies ist ein wichtiger Punkt, denn er besagt, dass durch strukturierte Interviews selbst keine neue Fragestellung für die Untersuchung entwickelt und nichts über die Bedeutung der gestellten Fragen für die Forschung erfahren werden kann. Was als wichtig angesehen wird und was nicht, ist durch die Konstruktion des Fragebogens bereits festgelegt.

Daraus folgt eine wichtige Konsequenz für den Einsatz von strukturierten Interviews und Fragebögen im Rahmen einer Feldforschung: Die Konstruktion eines sinnvollen Fragebogens, der im Rahmen einer übergeordneten Fragestellung eingesetzt wird, erfordert bereits erhebliches Wissen über die untersuchte Gruppe und das Thema, das erforscht werden soll. Eine standardisierte Befragung ist also keine Methode, mit der man sich einem neuen, weitgehend unbekannten Untersuchungsfeld nähern sollte. Dies gilt auch aus Gründen der Forschungspragmatik: Die Konstruktion eines Fragebogens und die Durchführung einer standardisierten Befragung sind so aufwendig, dass dafür vorhandene Ressourcen sehr überlegt und zielgenau eingesetzt werden sollten. Vorwissen ist dabei in mindestens zweierlei Hinsicht erforderlich: Einerseits muss schon bekannt sein, was tatsächlich relevante Fragen im Kontext der jeweiligen Forschung sind, andererseits muss man wissen, wie man die gewünschten Informationen erhalten kann, d. h. wie die beabsichtigten Fragen gestellt werden müssen. Eine Ausnahme stellt möglicherweise eine Zensuserhebung dar, bei der die Kategorien weitgehend von den Forschenden festgelegt werden und die manchmal auch als Einstieg in ein Forschungsfeld genutzt wird (siehe dazu den Beitrag von Julia Pauli in diesem Band).

6.4 Die Konstruktion von Fragebögen

Die Konstruktion eines Fragebogens sowie die Art des notwendigen Vorwissens hängen davon ab, welche Art von Informationen mit ihm erhoben werden sollen. Geht es etwa um allgemeine Daten, die eine Bevölkerung beschreiben, wie Alter, Geschlecht, Wohnort, Familienstand, Zahl der

Kinder usw.? Viele solcher Fragen benötigen eher geringes Vorwissen, da es um eine allgemeine Beschreibung der untersuchten Gruppe geht, deren Dimensionen weitgehend von dem/der Untersuchenden und seinem/ihrer Interesse festgelegt werden und in die kaum kulturelle Aspekte hineinspielen. Aber auch hier benötigt man einen gewissen Forschungsstand, um zu entscheiden, welche dieser Informationen zur Beantwortung übergeordneter Fragen erforderlich sind und welche nicht. Anders ist es bei Befragungen, die auf Wissen über Einstellungen, auf Werte oder kulturelles Wissen, also auf **Kognitionen**, zielen, sowie bei solchen, die nach bestimmten **Handlungen** fragen. Hier kann das erforderliche Vorwissen sehr umfangreich sein.

6.4.1 Der Fragebogen als Produkt vorhergehender Forschung

Wie erlangt man das Wissen, das man benötigt, um einen sinnvollen Fragebogen zu erstellen? In der Regel geht der Konstruktion eines Fragebogens eine längere Phase der Feldforschung voran, in der man erfährt, welche Aspekte eines Themas wichtig sind, welchen Aspekten von den Angehörigen der untersuchten Gruppe Bedeutung beigemessen wird, wie die Untersuchten über diese Aspekte sprechen (d. h. welche Terminologie und Kategorien sie verwenden) und über welche Aspekte es sinnvoll wäre, quantitative Daten zu erheben. Dieses Wissen kann durch unterschiedliche Vorgehensweisen erhoben werden: durch Teilnehmende Beobachtung, durch unstrukturierte Interviews, aber auch durch spezielle Verfahren wie *free-listing* (Fleisher und Harrington 1998), bei denen die Kategorien der InformantInnen zu bestimmten Themen gesammelt werden. Auf diese Art hat Susan Weller Krankheitskategorien bei Frauen in Guatemala und den USA erhoben (Weller 1984).

Es kommt also zunächst darauf an, Fragen zu entwickeln, die in den Fragebogen aufgenommen werden sollen. Diese Fragen werden auch **Items** genannt. Da man auf jede Frage verschiedene Antworten geben können muss (sonst wäre die Frage überflüssig), nennt man die erfragte Information Variable. Das Item „Was ist Ihr Geschlecht“ fragt nach der Variablen „Geschlecht“ und kann beispielsweise den Zustand „weiblich“ oder „männlich“ annehmen. Die meisten Items sind jedoch komplexer und können mehr als nur zwei oder drei Zustände annehmen.

6.4.2 Wie viele Fragen? Pragmatische Aspekte

Zunächst erstellt man also eine Liste der Items, die in einen Fragebogen aufgenommen werden sollen. Dabei ist zu beachten, dass die Beantwortung von Fragebögen Zeit kostet. Ausführliche Fragebögen mit vielen Items kosten

viel Zeit, und man kann davon ausgehen, dass das Interesse und die Bereitschaft der meisten InformantInnen, an einer Befragung teilzunehmen, mit dem erforderlichen Zeitaufwand abnehmen. Susan Weller (1998: 376) rät, lieber zu viele als zu wenige Fragen zu stellen, da es in der Regel unmöglich ist, später noch einmal zu allen InformantInnen zu gehen und die Fragen zu stellen, die zwar nicht im Fragebogen enthalten waren, die sich aber im Nachhinein als auch noch wichtig entpuppten. Ich halte das nur dann für einen sinnvollen Ratschlag, wenn ‚zu viele Fragen‘ nicht einen zu langen Fragebogen ergeben, den niemand mehr beantworten will. Dabei geht es natürlich nicht nur um den Zeitaufwand für die InformantInnen, sondern auch für diejenigen, die die Fragen stellen. Wenn Ressourcen für mehrere InterviewerInnen zur Verfügung stehen, kann in das einzelne Interview natürlich mehr Zeit investiert werden, als wenn eine Person sämtliche Interviews führen muss – und daneben auch noch andere Methoden anwenden möchte. Die Zeit, die ein Interview tatsächlich erfordert, ist nicht unbedingt im Voraus absehbar. Sie hängt nicht nur von der Länge des Fragebogens ab, sondern auch davon, ob tatsächlich nur die einzelnen Items durchgegangen werden oder ob auch Nebenbemerkungen der InformantInnen aufgezeichnet werden sollen. Manche InformantInnen antworten knapp und präzise, andere erzählen zu jeder Frage kürzere oder längere Geschichten, die es durchaus auch wert sein können, als Zusatzinformationen festgehalten zu werden. Der Fragebogen zur bereits erwähnten Absolventenstudie war sehr lang und umfasste 109 Fragen. Die einzelnen Interviews dauerten zum Teil über zwei Stunden. Die Konstruktion eines Fragebogens hat immer auch einen pragmatischen Aspekt. Es gilt, einen Mittelweg zu finden zwischen dem Wunsch, möglichst viele Informationen zu gewinnen, und der Möglichkeit, die dazu erforderlichen vielen Fragen auch wirklich zu stellen und beantwortet zu bekommen.

Aber natürlich hat Weller Recht, dass es ein schwerer Rückschlag für eine Erhebung ist, wenn man nach ihrer Durchführung feststellt, dass eine wichtige Frage nicht gestellt worden ist. Daraus folgt für mich aber weniger die Forderung, gleich so viele Fragen wie irgend möglich zu stellen, als die, sich über Sinn und Zweck der Erhebung im Rahmen der Gesamtforschung sowie über die Gestaltung eines Fragebogens sehr genau Gedanken zu machen und den Fragebogen gut zu testen. Ich denke aber auch, dass es sich zu einem gewissen Grad nicht vermeiden lässt, dass später noch Fragen auftauchen, die man gerne gestellt hätte. Da Forschung ein offener Prozess ist, der zu neuen Erkenntnissen führen soll, werden ständig neue Fragen aufkommen, die dem oder der Forschenden zu Beginn nicht in den Sinn gekommen sind. Das gilt besonders für ethnologische Feldforschung, die in der Regel immer auch einen explorativen Charakter hat und somit Aspekte umfasst, die im Voraus nicht genau gesehen werden können. Auch das ist ein Argument

dafür, eine standardisierte Befragung erst dann durchzuführen, wenn die Forschung schon fortgeschritten ist.

6.4.3 Die Formulierung von Fragen

Viele Fragen erscheinen zunächst einfach, sind es bei genauerem Hinsehen aber nicht. Wenn man beispielsweise türkische MigrantInnen nach ihrem Geburtsort fragt, bekommt man in der Regel nicht wirklich den Ort der Geburt genannt (d. h. ein Dorf oder eine Stadt), sondern die Provinz, in der dieser Ort liegt. Wenn man genauere örtliche Angaben haben will als die Provinz der Geburt, muss man also die Frage spezifizieren oder erklären, etwa durch den Zusatz, dass nicht (nur) die Provinz gemeint ist, sondern der tatsächliche Ort der Geburt. Welche Fragen einfach sind und welche nicht bzw. wie eine Frage formuliert werden muss, damit man tatsächlich das erfährt, was man wissen will, lernt man erst, wenn man die Fragen tatsächlich stellt und ausprobiert.

6.4.4 Geschlossene Fragen und offene Fragen

Grundsätzlich können Fragen in einem Fragebogen auf zwei Arten gestellt werden: mit vorgegebenen Antwortmöglichkeiten (geschlossene Fragen) oder ohne vorgegebene Antworten (offene Fragen). Die Frage nach dem Geschlecht, bei der man etwa je ein Kästchen zum Ankreuzen für „weiblich“, „männlich“ und „divers“ vorgibt, ist eine geschlossene Frage. Die Frage nach dem Geburtsort, bei der keine Liste möglicher Orte genannt wird, sondern der oder die Befragte selbst den Ort nennen soll, ist eine offene Frage. Man kann jede Frage als offene oder als geschlossene Frage formulieren, aber die Pragmatik wird geschlossene Fragen für manche Bereiche ausschließen. So wäre es wenig sinnvoll, eine Liste möglicher Geburtsorte in der Türkei zum Ankreuzen in den Fragebogen einzufügen, da er dann den Umfang eines mittleren Telefonbuchs hätte. Offene Fragen sind auch dann ratsam, wenn die Antwortmöglichkeiten auf eine Frage nicht im Voraus bekannt sind. Wenn ich also davon ausgehe, dass es in einer Gesellschaft nur zwei Geschlechter gibt, dann kann ich problemlos die Frage nach dem Geschlecht geschlossen formulieren und die beiden Antwortmöglichkeiten vorgeben. Wenn ich aber nicht genau weiß, ob es nicht vielleicht weitere Geschlechter gibt, dann muss die Frage offen formuliert werden, da die Geschlossenheit weitere Antwortmöglichkeiten ausschließen würde.

Geschlossene Fragen setzen also voraus, dass alle möglichen Antworten auf die Frage bekannt sind. Die vorgegebenen Antwortmöglichkeiten müs-

sen erschöpfend sein, d. h., es darf keine Antwortmöglichkeit geben, die nicht erfasst wird, und sie müssen wechselseitig ausschließend sein, es muss also eindeutig klar sein, welche der angebotenen Antworten zutrifft. Diese zweite Bedingung bedeutet nicht, dass nur jeweils eine Antwort zutreffend sein darf, da je nach Frage auch Mehrfachantworten möglich sein können. Wenn ich MigrantInnen aus der Türkei nach ihrer Staatsangehörigkeit frage und „deutsch“ und „türkisch“ als Antwortmöglichkeiten vorgebe, dann kann es sein, dass auf einige InformantInnen beide Antworten zutreffen, da sie zwei Staatsangehörigkeiten besitzen. Wichtig ist aber, dass die beiden Antwortmöglichkeiten klar voneinander abgegrenzt sind. Wenn etwa in der Türkei nach ethnischer Zugehörigkeit gefragt wird (u. a.) die beiden Antwortkategorien „Kurt“ (Kurde) und „Zaza“ vorgegeben werden, so sind diese beiden Antworten nicht klar voneinander abgegrenzt, da sehr oft – aber nicht immer – Zaza als eine Unterkategorie von Kurden verstanden wird. Für einen Zaza-Informanten wäre also nicht klar, ob er beide Antwortmöglichkeiten als auf sich zutreffend betrachten sollte oder nicht, und für den/die InterviewerIn ist nicht klar, ob der Informant „Zaza“ und „Kurt“ als exklusive Kategorien verstanden hat oder nicht.

6.4.5 Problematische Fragen

Manche Fragen werden schnell und ohne Vorbehalte beantwortet; anderen weicht man lieber aus. Oft sind die schwierigen Fragen gerade die, die uns besonders interessieren. Fragen nach Einkommen oder Besitz etwa, die der Feststellung von Schichtzugehörigkeit und ökonomischem Status dienen, wecken häufig Befürchtungen, die Steuerverwaltung könnte hinter der Befragung stehen oder sich zumindest die Ergebnisse nutzbar machen. Solche Befürchtungen sind nicht immer leicht zu zerstreuen, da man Fragebogenerhebungen ja bei vielen InformantInnen durchführt und gerade auch bei solchen, denen man nicht aus offenen Interviews oder Teilnehmender Beobachtung gut bekannt und vertraut ist. Das Hantieren mit Fragebögen hat immer einen gewissen ‚offiziellen‘ Charakter. Viele InformantInnen kennen Fragebögen vielleicht nur aus staatlichen Zensuserhebungen, bei denen der Zusammenhang zwischen Wissen und Macht bzw. staatlicher Kontrolle unmittelbar auf der Hand liegt. Man denke nur an die Erregung, die die Volkszählungspläne 1987 in Deutschland hervorriefen. Oft haben solche Erhebungen unmittelbare politische Relevanz und sind sehr konfliktträchtig. Gerade in Ländern mit weniger entwickeltem Datenschutz kann das Misstrauen groß und berechtigt sein. Für die Ergebnisse der Erhebung kann sich Misstrauen verheerend auswirken, weil es in der Regel zu kreativen Verweigerungsstrategien führt, nämlich zu falschen oder irreführenden Angaben.

Dies ist für das Ergebnis schlimmer als die direkte Weigerung zu antworten. Nicht gegebene Antworten kann man bei der Auswertung als solche behandeln. Falsche Angaben sind dagegen häufig nicht zu identifizieren. Man kann versuchen, ein gewisses Vertrauensverhältnis zu unbekannten InformantInnen aufzubauen, indem man sich über geeignete Mittelspersonen einführen lässt. Das wird aber bei einer Zufallsstichprobe in den meisten Fällen unmöglich sein.

Andere Fragen werden vielleicht gerade dann nicht gerne beantwortet, wenn der oder die InterviewerIn gut bekannt ist. Diese betreffen vor allem intime Lebensbereiche wie Sexualität. Fragen dazu werden manchmal als Bedrohung empfunden. Hierbei ist zu beachten, dass es natürlich kulturell unterschiedlich sein kann, welche Fragen Intimes berühren und damit problematisch sind und welche nicht. Man kann solche Fragen mit einem Satz einleiten, um ihnen etwas von ihrem ‚Bedrohungscharakter‘ zu nehmen. Ich halte jedoch nichts davon, derartige Fragen psychologisch besonders geschickt zu formulieren, um den InformantInnen eine Antwort zu entlocken, die sie eigentlich nicht geben wollen. Aus Gründen der Forschungsethik finde ich es besser, sich im Zweifelsfall mit „keine Antwort“ zufrieden zu geben.

In jedem Fall setzen beide Kategorien problematischer Fragen voraus, dass den InformantInnen verlässlich und nachvollziehbar Anonymität und Datenschutz zugesichert werden. Häufig ist es sinnvoll, dies in einer schriftlichen Erklärung zu tun – bei verschickten Fragebögen etwa in einem Begleitschreiben zum Fragebogen. Bei selbst ausgefüllten Fragebögen erfordert Anonymität, dass alle Angaben, die den oder die InformantIn identifizierbar machen, vom Hauptteil der Fragen klar getrennt werden.

Schließlich gibt es andere Fragen, die problematisch sind, obwohl sie weder Intimes noch steuerlich Relevantes berühren. Sie betreffen Dinge, deren Quantität nicht genau bekannt ist. So kann in vielen Gesellschaften die Frage nach dem Alter schwierig sein, da die meisten Menschen nicht genau wissen, wann sie geboren wurden. Auch ein Blick in den Ausweis, falls vorhanden, hilft hier nicht unbedingt weiter, da auch er oftmals nur ein geschätztes Geburtsdatum enthält. Früher war es beispielsweise in der Türkei üblich, neugeborene Kinder nur alle paar Jahre (und oft eben längere Zeit nach der Geburt) bei den Behörden anzumelden, mit der Konsequenz, dass die offiziellen Daten nicht verlässlich sind. Ähnliche Probleme können sich beim Landbesitz ergeben. Auch hier sind häufig keine genauen Flächenmaße bekannt. Man kann dann versuchen, sich mit Ersatzmaßen zu begnügen, etwa mit der Arbeitszeit, die benötigt wird, um die Felder zu pflügen, oder mit der Menge des Saatgutes, das für die Aussaat erforderlich ist.

6.4.6 Übersetzungen und Pretests

Häufig führen EthnologInnen ihre Befragungen nicht in ihrer eigenen Muttersprache durch. Fragebögen und andere Erhebungsinstrumente werden meistens in der Sprache des/der EthnologIn erstellt und dann in die Sprache der Gesprächspartner übersetzt. Uns ist sehr bewusst, dass sich Konzepte nicht in jedem Fall direkt aus einer Sprache in eine andere übersetzen lassen. Oft gibt es in der anderen Sprache kein Wort, das einen bestimmten Begriff aus der Sprache des Forschenden direkt wiedergibt. Homonyme oder Synonyme sind in den beiden Sprachen unterschiedlich, sodass auch dort, wo es direkte Entsprechungen gibt, das jeweilige semantische Feld differieren kann. Häufig gibt es aber auch die zentralen Kategorien der Fragen nicht in der Sprache der Ethnologen, wenn es etwa um lokale, emische Konzepte geht. All das erfordert große Sorgfalt bei der Übersetzung von Fragebögen. EthnologInnen sollten sich daher nicht auf die eigenen Kenntnisse der Sprache der InformantInnen verlassen, sondern die Übersetzung von MuttersprachlerInnen oder sogar von ausgebildeten ÜbersetzerInnen machen lassen. Manchmal ist es sinnvoll, mehrere Übersetzungen in Auftrag zu geben und miteinander abzugleichen. Viele AutorInnen schlagen vor, dass die Übersetzung wieder in die Ausgangssprache zurückübersetzt wird, um eventuelle Bedeutungsverschiebungen erkennen zu können. Brislin (1986) fordert sogar einen doppelten Durchgang durch diesen Übersetzungskreislauf, um Probleme ausschließen zu können.

Die Schwierigkeit der Auswahl und Formulierung von Fragen sowie der Übersetzung macht es erforderlich, dass sich nicht nur die Forschenden selbst kritische Gedanken über den Fragebogen macht, sondern dass man ihn tatsächlich mit mehreren Gesprächspartnern testet, bevor man mit der eigentlichen Untersuchung beginnt. Erst wenn die Fragen tatsächlich InformantInnen gestellt werden, wird deutlich, wo Bedeutungen unklar sind, wo Fragen missverstanden werden können, wo Einheiten problematisch sind oder sonstige Schwierigkeiten auftauchen. Erst bei einem **Pretest** wird etwa offensichtlich, welche impliziten Vorannahmen in den Fragen stecken, die von den InformantInnen nicht geteilt werden. Nach dem Pretest muss der Fragebogen auf der Basis der gewonnenen Erfahrungen gründlich überarbeitet werden. Wichtig ist, dass die Testpersonen nicht aus dem Sample stammen dürfen, das letztendlich befragt werden soll.

6.5 Sampling

6.5.1 Wer soll befragt werden?

Will man eine Erhebung in einer kleinen Siedlung oder in einer Nomaden-Gruppe durchführen, die nur wenige Haushalte umfasst, dann kann man tatsächlich alle Mitglieder der Gruppe befragen. Sehr oft, etwa bei Forschungen, die in Städten durchgeführt werden, ist das jedoch nicht der Fall. Man kann nur einen Teil der Gruppenmitglieder interviewen. In der Regel möchte man mit der Befragung aber zu Daten kommen, die nicht nur etwas über die tatsächlich Befragten aussagen, sondern über die gesamte Gruppe, aus der sie ausgewählt wurden. Dies ist deswegen problematisch, weil man davon ausgehen kann, dass die Gruppe nicht homogen ist. Wenn man Gruppenmitglieder A und B befragt hat, kann man daraus nicht schließen, dass C und D dieselben Antworten geben würden. Und wahrscheinlich haben schon A und B unterschiedliche Antworten gegeben. Das Zauberwort heißt hier **Repräsentativität**: Die Stichprobe von Gruppenmitgliedern (das *Sample*), die befragt werden, soll repräsentativ für die gesamte Gruppe sein, sodass Aussagen, die über das jeweilige Sample gemacht werden, mit großer Wahrscheinlichkeit auf die gesamte Gruppe zutreffen.

Sampling ist in der Regel eine schwierige Forschungsoperation. Eigentlich müsste man ja bereits die Gesamtgruppe kennen, um eine Stichprobe mit beispielsweise der gleichen Alters-, Geschlechts-, Einkommensstruktur usw. ziehen zu können, wie sie in der Gesamtgruppe vorherrscht. Da das nicht der Fall ist (denn dann wäre die Erhebung ja überflüssig), behilft man sich mit einem Zufallssample. ‚Zufall‘ heißt hier nicht, dass die InformantInnen einfach zufällig ausgewählt werden, denn ‚zufällig‘ in diesem Sinne wäre gleichbedeutend mit ‚willkürlich‘ oder ‚beliebig‘. Zufall bedeutet, dass jedes Mitglied der Gesamtgruppe dieselbe Chance haben muss, ausgewählt werden zu können. Es darf also keine Faktoren geben, die bewirken, dass bei einigen Gruppenmitgliedern eine größere Wahrscheinlichkeit besteht, in das Sample aufgenommen zu werden, als bei anderen.

Um ein Zufallssample herstellen zu können, muss man durchaus etwas über die Gesamtgruppe wissen: Man muss wissen, wer alles zur Gruppe gehört und in das Sample fallen kann. Man braucht einen Rahmen, ein *Sampling Frame*, aus dem gewählt wird. Ein Sampling Frame ist eine Liste, in der alle Mitglieder einer Gruppe verzeichnet sind. Das kann ein Einwohnerregister oder ein Telefonbuch (Bernard 2011: 115) sein, wenn es um eine Studie in einer Stadt oder einem Stadtteil geht, oder etwa die Liste der SchülerInnen einer Schule, wenn diese Schülerschaft untersucht werden soll. Falls kein Sampling Frame vorhanden oder zugänglich ist, muss man selbst einen Rahmen herstellen, etwa, indem man einen Zensus aufnimmt. Da dies bei

einer größeren Gruppe nahezu unmöglich ist, wird man in diesem Fall alle Kreativität daransetzen, doch ein bereits bestehendes Sampling Frame zu verwenden.

Nur wenn man einen Rahmen für das Sampling hat, kann man überlegen, wie man vorgehen muss, um daraus eine Zufallsstichprobe auszuwählen. Da jedes Individuum des Sampling Frames dieselbe Chance haben muss, ausgewählt werden zu können, hat es sich eingebürgert, mit Zufallszahlen vorzugehen, die mit vielen Statistikprogrammen generiert werden können. Man nummeriert alle Individuen des Rahmens durch und muss festlegen, wie viele Individuen im Sample enthalten sein müssen (siehe unten). Dann geht man die Zufallszahlen von einem beliebigen Einstieg aus der Reihe nach durch und wählt immer ein Individuum aus, wenn dessen Nummer in der Folge der Zufallszahlen vorkommt. Das macht man so lange, bis man so viele Individuen ausgewählt hat, wie das Sample enthalten soll.

Diese Methode ist bei sehr großen Sampling Frames impraktikabel, da man nicht alle Mitglieder einer großen Gruppe durchnummerieren kann. Man kann sich dann mit einem **Sampling-Intervall** behelfen. D. h., man beginnt mit einem zufällig ausgewählten Individuum in der Liste des Sampling Frame und wählt jedes n -te Individuum in der Folge aus, wobei n für das Intervall, den „Abstand“ zwischen den jeweils ausgewählten Individuen steht. Wenn man bei einer Gruppe von 10.000 Individuen jedes 25. Individuum nimmt, bekommt man ein Sample von 400 (n wäre hier also 25). Das Intervall ist demnach abhängig von der Größe des Rahmens und des gewünschten Samples. Beim Intervall-Sampling muss man darauf achten, dass das Sample nicht durch eine mögliche Regelmäßigkeit der Gesamtgruppe beeinflusst wird. Angenommen, man will in einer großen Reihenhaussiedlung jeden zehnten Haushalt befragen, und die Siedlung ist so aufgebaut, dass jedes zehnte Haus ein Eckhaus mit großem Garten ist. Dann kann es passieren, dass das Sample entweder nur oder gar keine Eckhäuser enthält. Das Sample wäre also nicht repräsentativ. Wenn eine solche Periodizität erkennbar ist, muss man ein Sample-Intervall wählen, das nicht parallel zu dieser Regelmäßigkeit verläuft.

Falls man ein Sample aus einer Gruppe ziehen will, die nach einem bestimmten Kriterium (z. B. Migrationserfahrung) eine zahlenmäßig nicht sehr große, für Vergleichszwecke der Untersuchung aber wichtige Subpopulation enthält, dann kann es sinnvoll sein, für jede der Gruppen ein eigenes Zufallssample zu bilden. Angenommen, zehn Prozent einer Gesamtgruppe von 10.000 sind MigrantInnen und man möchte ein Sample von 400 bilden, dann kann man bei der Zufallsauswahl eines Gesamtsamples nicht unbedingt davon ausgehen, dass man darunter gerade 40 MigrantInnen (= zehn Prozent) auswählt. Vielleicht enthält das Zufallssample nur 30 MigrantInnen. In diesem Fall ist es besser, gleich zwei Samples zu bilden: eins mit

360 Nicht-MigrantInnen und eins mit 40 MigrantInnen. In diesem Fall hat man zwei proportionale Samples gebildet. Wenn eine Untergruppe, die man erfassen möchte, relativ zur Gesamtbevölkerung sehr klein ist, kann es auch sinnvoll sein, nicht-proportionale Samples zu bilden. Man wählt dann im Vergleich zur Hauptgruppe überproportional viele Angehörige der Untergruppe aus, um zu gewährleisten, dass die Untergruppe im Sample ausreichend vertreten ist.

Falls kein praktikabler Sampling Frame zur Verfügung steht, da die zu untersuchende Gruppe zu groß und zu unübersichtlich ist, kann man sich mit **Cluster Samples** behelfen. Bei einer Stadtforschung unterteilt man etwa die Stadt in Stadtteile, Nachbarschaften oder Häuserblocks (die Cluster), unter denen man ein Zufallssample auswählt. In diesen zufällig ausgewählten Clustern kann dann jeweils ein repräsentatives Sample erstellt werden.

6.5.2 Wie groß soll ein Sample sein?

Wir wissen nun, wie man Samples auswählen kann, aber das beantwortet noch nicht die Frage danach, wie groß ein Sample zum Zweck der Untersuchung sein soll. Die notwendige Größe des Samples hängt mit seiner Heterogenität zusammen. Wenn alle Individuen in einer Stadt mit einer Million Einwohner völlig identisch wären, dann bräuchte man nur eine Person zu interviewen, um Informationen über alle Bewohner zu haben. Derartige extreme Homogenität ist aber beim Untersuchungsgegenstand der Ethnologie selbstverständlich nicht vorhanden. Da im Voraus nicht bekannt ist, wie homogen oder heterogen eine Gruppe ist, geht man auf Nummer sicher, indem man ihre größtmögliche Heterogenität annimmt. Die Größe des erforderlichen Samples wächst mit der Größe der Gesamtgruppe, über die etwas ausgesagt werden soll. Allerdings wächst die Größe des Samples nicht proportional zur Gesamtgruppe. Im Gegenteil: Für kleine Gesamtgruppen ist ein relativ großes Sample erforderlich, für große Gesamtgruppen aber nur ein relativ kleines Sample. Die Größe des erforderlichen Samples hängt dabei auch von der gewünschten Wahrscheinlichkeit ab, dass das Sample tatsächlich die Gesamtgruppe repräsentiert, d. h. dass die Verteilung eines bestimmten Merkmals im Sample tatsächlich der Verteilung desselben Merkmals in der Gesamtgruppe entspricht. Bernard gibt an, dass bei einer Gesamtgruppe von 50 Personen ein Sample von 44 erforderlich ist, um mit einer Wahrscheinlichkeit von 95 % der Gesamtgruppe zu entsprechen. Bei einer Gesamtgruppe von 1.000.000 wäre es hingegen ein Sample von 384 (1995: 79, Tabelle 4.3). Mathematisch Interessierte finden bei Bernard auch die Formel, mit der die Samplegröße berechnet werden kann (ebd.: 78).

Man kann sich den Zusammenhang zwischen Größe der Gesamtgruppe und Größe des repräsentativen Samples mit einem Gedankenexperiment verdeutlichen: Angenommen, man will Aussagen über eine ‚Population‘ von Würfelwürfen machen. Hat man nur eine kleine Anzahl von Würfeln, etwa zehn, dann ist die Wahrscheinlichkeit gering, dass alle möglichen Wurfresultate gleich häufig vertreten sind. Man braucht also möglichst Informationen über alle zehn Würfe, um diese Population adäquat beschreiben zu können. Die Größe des Samples muss also der Größe der Gesamtpopulation gleichen. Ist die Wurfpopulation aber sehr groß, dann ist auch die Wahrscheinlichkeit sehr groß, dass die Würfe sich einer Normalverteilung annähern, dass also alle Zahlen gleich häufig fallen (vorausgesetzt, der Würfel hat keinen Fehler). Diese Normalverteilung ändert sich nicht durch häufigeres Würfeln. Man kann davon ausgehen, dass bei hundert zufällig ausgewählten Würfeln die Zahlen nicht anders verteilt sind als bei einer Million Würfeln, obwohl das Sample nur einen Bruchteil der Gesamtpopulation von Würfeln umfasst.

6.6 Erhebungswege

Wie werden die Fragen den InformantInnen bzw. dem ausgewählten Sample nun tatsächlich gestellt? Grundsätzlich können über die verwendeten Kommunikationswege drei Arten der Befragung unterschieden werden: a) mündliche Befragung durch InterviewerInnen; b) telefonische Befragung durch InterviewerInnen; c) schriftliche Befragung, bei der die InformantInnen einen vorliegenden Fragebogen selbst ausfüllen. Für alle Erhebungsarten gilt, dass die Erhebung in einem möglichst kurzen Zeitraum durchgeführt werden sollte, um zu vermeiden, dass Ereignisse eintreten, die das Antwortverhalten der später Interviewten entscheidend beeinflussen.

6.6.1 Direkte mündliche Befragung

In vielen Fällen ergibt sich die Erhebungsweise schon aus der Art und dem Ort der Forschung.

In vielen Gebieten, in denen EthnologInnen klassischerweise ihre Feldforschungen durchführen, sind andere Erhebungsweisen als die direkte mündliche Befragung durch eine/n Interviewenden nicht möglich. Eine telefonische Befragung ist oft schwierig, weil zwar inzwischen sehr viele Menschen über ein Mobiltelefon verfügen, es aber gerade wegen der Mobilität häufig schwierig ist, ein längeres, konzentriertes Gespräch für die

Befragung zu führen. Eine schriftliche Befragung ist vielleicht nicht möglich, weil viele der InformantInnen des Lesens und Schreibens nicht kundig sind.

Mündliche Interviews sind in der Regel zeitaufwendig. Das gilt nicht nur, wenn der Fragebogen sehr umfangreich ist, sondern auch weil man in der Regel nicht mit einem Fragebogen ‚ins Haus fallen‘ und gleich wieder verschwinden kann, wenn denn alle Fragen abgehakt sind. Die mündliche Befragung ist in der Regel in einen ‚Besuch‘ eingebettet, und ein ‚Besuch‘ kann je nach kulturellem Kontext ein relativ langandauerndes Ereignis sein. Bevor man zum Fragebogen kommt, muss man vielleicht verschiedene Begrüßungsrituale absolvieren und ausführlich ‚small talk‘ führen – wobei ‚small‘ hier ein sehr irreführendes Adjektiv sein kann. Oft erfordert ein Besuch, dass der Besuchte, der sich dann als Gastgeber versteht, dem Besucher und Gast etwas zu essen oder zu trinken anbietet, was der oder die Interviewende als Gast wiederum nicht ablehnen darf. Es kann also länger dauern, bis man überhaupt dazu kommt, den Fragebogen zu zücken. Und wenn er denn abgearbeitet ist, müssen vielleicht analog zur Ankunft diverse Abschiedsformalitäten eingehalten werden. Diese ‚kommunikative Einbettung‘ einer Befragung muss natürlich kein Nachteil sein, da man auch beim Gespräch vor und nach dem Interview zahlreiche Informationen erhalten kann.

In einem solchen Kontext kann man also nicht davon ausgehen, dass ein Interview nur so viel Zeit erfordert, wie man etwa für das testweise Durchgehen der Fragen mit ProbeinformantInnen benötigt. Unter Umständen kann man sich glücklich schätzen, wenn man pro Tag zwei Interviews durchführen kann. Wenn man für die Feldforschung ein Jahr Zeit hat, aber nicht das ganze Jahr mit standardisierten Interviews verbringen möchte, kann man sich ausrechnen, dass man als einzelner Forscher auf diese Art nicht sehr viele Personen befragen kann.

Hier ist der Einsatz von ForschungsassistentInnen sehr sinnvoll. Wenn man einen Antrag auf Forschungsmittel stellt, sollte man also gleich überlegen, welche Befragungen geplant sind und ob es sinnvoll ist, Mittel für AssistentInnen zu beantragen. Wenn man AssistentInnen engagiert, muss man sie jedoch sehr gründlich auf ihre Aufgabe vorbereiten. Sie müssen beispielsweise die Fragen so gut verstehen, dass sie sie auf Nachfrage den Interviewten erläutern können. Auch dürfen sie die Frage nicht auf eine suggestive Art und Weise stellen, die den InformantInnen schon eine Antwort in den Mund legt. In der Regel ist es sinnvoll, dass man nach einer theoretischen Anleitung mehrere Interviews gemeinsam mit den AssistentInnen durchführt, ihr Frageverhalten beobachtet und nach der Interviewsituation mit ihnen bespricht und korrigiert. Axinn et al. (1991) gehen ausführlicher auf das Training von InterviewerInnen ein.

6.6.2 Telefonische Befragung

Telefonische Befragungen assoziiert man eher mit Wahl- oder Marktforschung, aber sie kann auch in der Ethnologie sinnvoll sein. Das ist der Fall, wenn einerseits Telefone weit genug verbreitet sind, sodass man seine InformantInnen tatsächlich telefonisch erreichen kann, und wenn andererseits die InformantInnen so weit verstreut leben, dass man sie gar nicht alle zu einem Interview aufsuchen kann. Unsere Befragung der AbsolventInnen des Hamburger Ethnologiestudiums haben wir aus genau diesem Grund telefonisch durchgeführt – es lebten eben nicht mehr alle AbsolventInnen in Hamburg oder auch nur in Deutschland.

Telefonische Interviews lassen sich oft schneller erledigen als Interviews im direkten Gespräch, da man am Telefon häufig direkter ‚zur Sache‘ kommen kann und die sonst vielleicht obligatorische Tasse Kaffee beim Gespräch ausfällt. Gleichzeitig ist die Kommunikationsatmosphäre häufig unverbindlicher. Das kann eine gewisse Anonymität herstellen, in der sich manche Dinge leichter fragen und beantworten lassen als im persönlichen Gespräch. Genauso wie im persönlichen Gespräch sind auch beim Telefoninterview Nachfragen durch die InformantInnen möglich, Unklarheiten beim Interview können also ausgeräumt werden. Außerdem geben die GesprächspartnerInnen vielleicht ‚nebenher‘ Informationen, die nicht im Fragebogen abgefragt werden, aber trotzdem nützlich sind. Ein Nachteil ist, dass man beim Telefoninterview die Mimik und Gestik der GesprächspartnerInnen nicht sieht, die oft wichtige Hinweise zur Interpretation des Gesagten geben. Eine Schwierigkeit besteht darin, überhaupt die Telefonnummern herauszufinden. So konnten wir bei unserer Absolventenbefragung nur von gut der Hälfte der AbsolventInnen die aktuellen Telefonnummern ermitteln. Inzwischen ist es noch schwieriger geworden, Telefonnummern zu ermitteln, da viele Personen keinen Eintrag im Telefonbuch haben.

6.6.3 Schriftliche Befragung

Bei der schriftlichen Befragung wird den Befragten der Fragebogen übergeben, den sie dann selbstständig, ohne Beteiligung eines/einer Interviewenden ausfüllen. Das hat zunächst einige Nachteile. Nachfragen sind nicht möglich, der oder die Interviewte interpretiert den Fragebogen auf seine/ihre Art, und diese Interpretation kann von denen, die die Fragebögen erstellt und verteilt haben, nicht nachvollzogen werden. Man weiß also nicht, ob Befragte die Fragen ‚richtig‘ verstanden haben. Die Befragten haben ja keine weiteren Informationen über die Befragung als die, die ihnen schriftlich vorliegen, und umgekehrt können Forschende nicht aus Neben-

informationen, wie sie im Gespräch selbstverständlich sind, auf die Art des Verständnisses schließen. Diese Schwierigkeit erfordert besondere Sorgfalt bei der Erstellung von Fragebögen. Fragen müssen einfach und eindeutig formuliert werden. Ein schriftlicher Fragebogen muss außerdem von einem Anschreiben begleitet werden, das Sinn und Zweck der Befragung erläutert, Datenschutz und Anonymität zusichert und auch eine Kontaktadresse für die InformantInnen nennt, falls sie doch Rückfragen haben.

Schriftliche Befragungen sind in der Marktforschung sehr üblich. In der Ethnologie können sie dann sinnvoll sein, wenn man ein großes Sample befragen möchte und dazu die Möglichkeiten direkter oder telefonischer Befragungen nicht ausreichen. Eine schriftliche Befragung erfordert eine Form der Kontrolle darüber, wer einen Fragebogen bekommt und wer nicht. Fragebögen einfach per ‚Schneeballsystem‘ zu verteilen, ist nicht sinnvoll, da dann keine Kontrolle des Samples – und somit keine Repräsentativität – möglich ist.

Während früher Fragebögen häufig per Post versandt wurden, finden schriftliche Befragungen heute in der Regel über das Internet statt. Es gibt verschiedene Online-Werkzeuge für Befragungen, wie zum Beispiel *SoSci Survey* oder *SurveyMonkey*, die für nicht-profitorientierte Zwecke oder bei einem kurzen Fragebogen teilweise kostenlos genutzt werden können. Die Online-Fragebögen sind in der Konstruktion häufig relativ einfach und nutzerfreundlich. Die Nutzerfreundlichkeit bei der Erstellung hat jedoch oft zur Folge, dass schnell sehr große und verschachtelte Fragebögen entstehen, die komplexe Datensätze generieren, deren Auswertung ebenfalls sehr komplex ist. Auch hier gilt, dass man sich vorher sehr genau darüber Gedanken machen sollte, welche Daten man tatsächlich benötigt, und bemüht sein sollte, den Fragebogen möglichst einfach zu halten. Um die Kontrolle des Samples zu gewährleisten, werden Fragebögen in der Regel nicht allgemein zugänglich ins Netz gestellt. Stattdessen wird der Link des Fragebogens bzw. ein Code, mit dem der Link freigeschaltet werden kann, per Email an das zuvor ausgewählte Sample verschickt. Damit kann man auch kontrollieren, wer den Fragebogen bereits ausgefüllt hat (auch wenn die Daten selbst anonym erhoben werden), und eventuell Erinnerungsmails an Personen verschicken, die das noch nicht getan haben. Der Versand von Erinnerungen erhöht in der Regel die Antwortquote massiv.

6.7 Auswertung

Wir haben nun also einen Fragebogen erstellt, ein Sample ausgewählt und die Befragung durchgeführt. Die ausgefüllten Fragebögen stapeln sich im

Zelt neben dem Schlafsack des Forschers oder auf dem heimischen Schreibtisch. Die Fragebögen enthalten Daten. Was fängt man damit nun an?

Das Thema dieses Bandes ist die Datenerhebung. Verfahren zur Datenauswertung würden ein eigenes Buch füllen. Da es aber unbefriedigend wäre, auf die Auswertung der erhobenen Daten überhaupt nicht einzugehen, sollen hier wenigstens einige einführende Informationen gegeben werden. Die Auswertung der erhobenen Daten geschieht heute in der Regel mit entsprechenden Computerprogrammen, wie z. B. dem *Statistical Package for the Social Sciences* (SPSS) oder der Open-Source-Software *R*. Viele statistische Operationen lassen sich auch mit Standard-Tabellenprogrammen wie Excel durchführen. Das befreit uns davon, selbst komplizierte Rechnungen durchführen zu müssen. Die zielführende Benutzung derartiger Programme setzt aber Statistikkennntnisse voraus, damit man weiß, welche Operationen mit welchen Daten sinnvoll durchgeführt und wie diese dann inhaltlich interpretiert werden können. An dieser Stelle kann ich keinen Crashkurs in Statistik geben, nur einige wichtige Basiskennntnisse sollen vermittelt werden. Eine Einführung in die Statistik, die gleichzeitig die Arbeit mit SPSS vorstellt, bieten Toutenburg und Heumann (2009).

6.7.1 Typen von Variablen

Mit den Fragen eines Fragebogens werden verschiedene **Variablen** erhoben. Diese Variablen können sehr unterschiedlichen Charakter haben. Man unterscheidet drei Messniveaus von Variablen: nominal, ordinal und metrisch. Nominale Variablen erfassen Eigenschaften, die an sich keine Quantitäten wiedergeben. Geschlecht ist eine nominale Variable, und männlich ist nicht mehr als weiblich. Nominale Variablen geben Eigenschaften wieder und sind dadurch charakterisiert, dass man lediglich die Häufigkeit ihrer Zustände feststellen, aber keine weitergehenden Rechenoperationen mit ihnen durchführen kann. So kann man auszählen, wie viele InformantInnen ledig oder verwitwet sind, es ist aber unmöglich, beispielsweise einen Mittelwert zwischen den verschiedenen Eigenschaften anzugeben. Das gleiche gilt für ordinale Variablen, die Rangfolgen erfassen, wie etwa den Schulabschluss oder den sozialen Status. Hier ist zwar ein ‚mehr‘ oder ‚weniger‘ impliziert, das Abitur gilt eben ‚mehr‘ als ein Hauptschulabschluss. Aber auch hier können weder Mittelwerte gebildet, noch der Abstand zwischen zwei Merkmalsausprägungen quantifiziert werden.

Dies ist bei numerischen Variablen anders, die tatsächlich Quantitäten angeben. Das Alter in Jahren oder die Haushaltsgröße sind numerische Variablen, die keine Eigenschaften erfassen, sondern Mengen und daher weiteren Rechenoperationen zugänglich sind. Numerische Variablen können eine dis-

krete Ausprägung haben und nur ganzzahlige Ergebnisse liefern, wie etwa die Zahl der Kinder in einem Haushalt. Oder sie sind kontinuierlich und können im Prinzip jede Zahl annehmen, was beispielsweise für die Größe von Landbesitz oder das Gewicht der Ernte zutrifft.

6.7.2 Die Kodierung der Daten

Bevor quantitativ erhobene Daten ausgewertet werden können, muss man sie kodieren. Dabei wird jeder möglichen Merkmalsausprägung einer Variablen ein Code zugeordnet, in der Regel eine Zahl. Die Kodierung wird in einem Codebuch festgehalten, damit auch im Nachhinein die Bedeutung der Codes wieder expliziert und nachvollzogen werden kann. Für die Variable „Familienstand“ kann etwa festgelegt werden: ledig = 1; verheiratet = 2; geschieden = 3; verwitwet = 4. Die Zuordnung der Codes ist willkürlich. In der Datentabelle (siehe unten) werden dann nur noch die entsprechenden Codes notiert. Man darf sich dabei aber nicht durch die Tatsache verwirren lassen, dass auch Nominalvariablen mit Ziffern ausgedrückt werden. Mit diesen Ziffern darf nicht gerechnet werden! Man kann nicht „ledig“ (1) und „verwitwet“ (4) addieren und dann als Mittelwert des Familienstandes den Wert 2,5 errechnen. Um das zu verhindern, muss im Codebuch unbedingt festgehalten werden, welchem Typ die jeweilige Variable entspricht, d. h. ob es sich um eine nominale, ordinale oder numerische Variable handelt. Neben der exakten Beschreibung einer Variablen mit den möglichen Merkmalsausprägungen und dem Typus enthält ein Codebuch normalerweise auch die laufende Nummer der Variablen (die meistens der Nummer der Frage im Fragebogen entspricht) sowie eine Abkürzung als Namen der Variablen.

Im Abschnitt über die Konstruktion von Fragebögen wurde zwischen offenen und geschlossenen Fragen unterschieden. Geschlossene Fragen, bei denen die Antwortkategorien (die vollständig und disjunkt sein müssen) schon vorgegeben sind, erleichtern die Kodierung. Man kann gleich auf dem Fragebogen hinter jedem Kästchen zum Ankreuzen den zugeordneten Code angeben. Die Kodierung offener Fragen ist schwieriger. Sie kann erst nach der Befragung vorgenommen werden, da vorher gar nicht bekannt ist, welche Antworten vorkommen werden. In diesem Fall müssen aus den gegebenen Antworten sinnvolle Kategorien gebildet werden, denen man dann Codes zuordnet.

Numerische Daten müssen nicht notwendigerweise kodiert werden. Bei der Variable „Alter“ kann einfach das Alter als Anzahl der Lebensjahre in die Datentabelle eingetragen werden. Es ist aber unbedingt erforderlich, dass die Einheit im Codebuch festgehalten wird, also etwa, dass das Alter

in Jahren angegeben wird. Es kann auch sein, dass man numerische Daten nicht als solche weiterverarbeiten will, sondern dass sie in Gruppen zusammengefasst werden sollen. So ist es vielleicht sinnvoll, die Variable „Einkommen“ nicht in absoluten Zahlen festzuhalten, sondern in Tausendergruppen, etwa „bis 1000“, „1001 bis 2000“, „2001 bis 3000“ und so weiter. Dann wird jeder Gruppe ein Code zugeordnet. In solchen Fällen muss man sich überlegen, ob gleich bei der Datenerhebung nur danach gefragt werden soll, in welche Gruppe das Einkommen der InformantInnen fällt, oder ob das exakte Einkommen erhoben wird und man erst im Nachhinein Gruppen bildet. Dabei ist es wichtig zu beachten, dass Daten, die nur in Gruppen erhoben werden, nicht wieder in exakte Angaben zurückverwandelt werden können, auch wenn das zu einem späteren Zeitpunkt der Analyse sinnvoll wird. Man hält sich also mehr Möglichkeiten der Analyse offen, wenn Daten zunächst so exakt wie möglich erhoben und erst nachträglich in Gruppen zusammengefasst und kodiert werden.

Häufig liegen nicht von allen InformantInnen Daten zu allen Variablen vor, etwa, weil ein/e InformantIn eine Frage nicht beantworten kann oder will. Derartige fehlende Daten werden gesondert kodiert (oft mit „9999“), damit sie bei der Analyse als solche erkannt und behandelt werden können.

6.7.3. Die Datentabelle

Die kodierten Daten können nun in eine Datentabelle oder Matrix eingetragen werden. Die Erhebung der Daten mit einem Online-Fragebogen hat den großen Vorteil, dass die Daten in der Regel bereits als Tabelle ausgegeben werden. Bei allen anderen Erhebungsformen müssen die Daten per Hand eingegeben werden. Die Datentabelle ist so organisiert, dass sich alle Angaben zu einer/r InformantIn in einer Reihe befinden. Alle Daten, die zu einer Person erhoben wurden, bilden einen Datensatz. Wenn man von 100 InformantInnen Daten erhoben hat, liegen also 100 Datensätze vor. Die Variablen sind in Spalten angeordnet. Auch Statistikprogramme wie SPSS arbeiten mit solchen Datentabellen. Dabei müssen die Spalten jeweils für die entsprechende Variable eingerichtet werden. Dazu wird der Spalte der Name der Variable zugeordnet, man muss die Art der Variablen, ihre Kodierung und die Anzahl der Stellen angeben, die für die Aufnahme der Codes erforderlich sind. Die erste Spalte der Datenmatrix enthält normalerweise die laufende Nummer des Datensatzes (d. h. die Identifikationsnummer des Informanten), und daran schließen sich Spalten zu allen Variablen an. Eine erhobene Variable kann auch in mehreren Variablen in der Datentabelle ausgedrückt werden. So kann man das Einkommen sowohl in einer Spalte als

exakten Wert angeben und in einer zweiten Spalte dann den Code für die Gruppe nennen, in die das jeweilige Einkommen fällt.

Mit Statistikprogrammen ist es kein Problem, auch große Datentabellen zu verwalten. Wenn man 100 InformantInnen befragt hat und von jedem und jeder/r von ihnen 60 Variablen erhoben hat, dann umfasst die Matrix schon 6000 Felder.

Die Eingabe von Daten in eine Tabelle ist eine ermüdende Tätigkeit, besonders bei zahlreichen und großen Datensätzen. Umso wichtiger ist es, dass die Dateneingabe mit großer Sorgfalt geschieht, denn Fehler bei der Dateneingabe gehören zu den häufigsten Fehlerquellen bei der Auswertung. Es ist daher wichtig, die Daten nach der Eingabe noch einmal zu überprüfen. „Ausreißer“ und unplausible Werte können oft schnell erkannt werden, etwa wenn ein Einkommen oder Alter außergewöhnlich hoch ist. Man sollte die Daten auch untereinander auf ihre Konsistenz prüfen, ob z. B. die angegebene Schulbildung mit dem ausgeübten Beruf zusammenpasst. Natürlich sind auf den ersten Blick inkonsistent erscheinende Daten oder Ausreißer nicht in jedem Fall falsch, aber es lohnt sich, sie noch einmal extra zu überprüfen.

6.7.4. Univariate Statistik

Die quantitative Auswertung von Daten lässt sich in zwei große Bereiche unterteilen. Der erste betrifft die jeweils einzelnen Variablen und beschreibt die für sie erhobenen Daten. Da jede Variable für sich betrachtet wird, spricht man von univariater oder auch von deskriptiver Statistik. Der zweite Bereich betrifft die Analyse von Relationen zwischen verschiedenen Variablen, mit denen spezifische Hypothesen getestet werden. Ich gehe in diesem Beitrag lediglich auf die univariate Statistik ein.

Drei Bereiche univariater Statistik können unterschieden werden: Häufigkeitsverteilung, zentrale Tendenz und Streuung. Zur Feststellung der Häufigkeitsverteilung wird errechnet, wie oft jede Merkmalsausprägung einer Variablen im Sample vorkommt. Dazu legt man eine Häufigkeitstabelle an. Normalerweise wird die Häufigkeit jedes Zustandes der Variablen sowohl in absoluten Zahlen als auch in Prozenten angegeben. Oft werden zusätzlich zu den Angaben zu den einzelnen Merkmalsausprägungen auch jeweils die kumulierten Angaben (meist in Prozenten) errechnet. Kumulierte Angaben haben den Vorteil, dass man auf einen Blick die gemeinsame Häufigkeit mehrerer Merkmalsausprägungen sehen kann. So kann man in der folgenden Tabelle der Anzahl von Kindern in einem Haushalt auf einen Blick erkennen, dass in 65,7 % der Haushalte höchstens zwei Kinder leben. Konventionell wird zu der Tabelle die Größe des Samples (n) angegeben.

Häufigkeitstabelle: Zahl der Kinder im Haushalt

Zahl der Kinder	Häufigkeit	Prozent	kumulierte Prozente
0	5	14,3	14,3
1	7	20,0	34,3
2	11	31,4	65,7
3	8	22,9	88,6
4	3	8,6	97,1
5	1	2,9	100,0

n = 35

Um die Häufigkeitsverteilung anschaulich zu machen, bietet es sich an, sie nicht nur in einer Tabelle, sondern auch in einem Diagramm darzustellen (Balken- oder Kreisdiagramm).

Sogenannte Maße der zentralen Tendenz geben die „Mittelwerte“ an – wobei Mittelwert hier nicht unbedingt das arithmetische Mittel bedeutet. Das **arithmetische Mittel** ist sicherlich das bekannteste Maß der zentralen Tendenz. Es errechnet sich aus der Summe der erhobenen Werte geteilt durch die Anzahl der Fälle. Im Fall der Anzahl der Kinder im Haushalt ergibt sich so ein arithmetisches Mittel von zwei Kindern. Das arithmetische Mittel kann nur für numerische Variablen errechnet werden. Ein weiteres Maß ist der **Median**. Der Median unterteilt die erhobenen Werte so, dass er genau in der Mitte liegt, d. h., die eine Hälfte der aufgetretenen Werte liegt ‚vor‘ dem Median, die andere Hälfte folgt ‚nach‘ dem Median. Ein weiteres Maß der zentralen Tendenz ist der Modus oder Modalwert. Er ist der im Sample am häufigsten vorkommende Wert, d. h., es können bei einer Messung auch mehrere Modi vorkommen, wenn mehrere Werte gleich häufig gemessen werden.

Mit der Angabe von Mittelwerten allein ist ein Sample unvollständig beschrieben. Darauf weist ein alter Statistikerwitz hin: Ein Statistiker hat die rechte Hand im Eisfach und die linke auf einer heißen Herdplatte. Er kommentiert: ‚Im Mittelwert eine angenehme Temperatur!‘ Maße der **Streuung** liefern Informationen über die Verteilung der Werte in einem Sample. Hier ist zunächst die **Spannweite** zu nennen, die sich aus der Differenz zwischen dem größten und dem kleinsten gemessenen Wert ergibt. In die Spannweite gehen nur zwei der gemessenen Werte ein, die **Standardabweichung** be-

rücksichtigt dagegen alle Werte. Sie errechnet sich, indem man die Wurzel aus der durchschnittlichen quadratischen Abweichung vom Mittelwert bildet. Die Standardabweichung ist ein Maß für die Homogenität oder Heterogenität der Messwerte: Werden nur identische Werte gemessen, dann ist die Standardabweichung null.

Mit Häufigkeiten, Maßen der zentralen Tendenz und Maßen der Streuung lassen sich die Erhebungsergebnisse zu den einzelnen Variablen beschreiben. Wichtig ist, dass es hier immer nur um einzelne Variablen geht, nicht um Beziehungen und Zusammenhänge zwischen mehreren Variablen.

6.8 Schluss

Quantitative Methoden werden mehr mit Soziologie als mit Ethnologie assoziiert. Dass auch die Ethnologie ohne Quantifizierung nicht auskommt, hat dieser Beitrag gezeigt. In manchen Forschungssettings gelten Fragebögen geradezu als Ausweis sozialwissenschaftlicher Forschung: Ein Ethnologe, der ‚nur‘ offene Interviews mit seinen Gesprächspartnern führt, wird manchmal nicht als ‚richtiger Wissenschaftler‘ anerkannt. In Pakistan habe ich das mehrfach erlebt.

Die Verfahren, die Soziologie und Ethnologie bei quantitativen Erhebungen anwenden, sind identisch. Aber es gibt trotzdem Unterschiede: Da EthnologInnen häufiger als SoziologInnen in kulturell anderen Kontexten arbeiten, müssen sie bei der Erstellung von Fragebögen, bei der Auswahl von Kategorien und der Formulierung der Fragen noch größere Sorgfalt walten lassen als die KollegInnen aus der Soziologie. Und schließlich kommt in der Ethnologie etwas kaum vor, was in der Soziologie recht häufig ist: dass eine Forschung allein mit quantitativen Daten operiert. Man mag das – etwa aus der Perspektive der empirischen Sozialforschung – für einen Mangel an Exaktheit halten, tatsächlich drückt sich darin aber die Erkenntnis aus, dass sich nicht alle Aspekte menschlicher Lebenswelten sinnvoll in Zahlen erfassen lassen. Ethnologische Forschung wird daher nie ohne eine Kombination verschiedener Erhebungsmethoden auskommen.

6.9 Zur Weiterarbeit empfohlene Literatur

Das Handbuch von H. Russel Bernard, *Research Methods in Anthropology* (2011), ist weit mehr als eine Einführung in die Feldforschungsmethoden der Ethnologie. Es ist ein Handbuch, dass quantitative Ansätze ebenso detailliert ausführt wie qualitative Methoden.